

Matrix-Lifting Semi-Definite Programming for Decoding in Multiple Antenna Systems

Amin Mobasher and Amir K. Khandani

Coding & Signal Transmission Laboratory (www.cst.uwaterloo.ca),

Dept. of Elec. and Comp. Eng., University of Waterloo, Waterloo, ON, Canada, N2L 3G1,

E-mail: {amin, khandani}@cst.uwaterloo.ca

Abstract—Recently, several quasi-maximum likelihood decoding methods have been introduced to solve the decoding problem in multiple antenna systems. Mobasher et al. [1] proposed a general method with a near optimal performance for M-ary QAM or PSK constellation. However, it is more complex compared to some other methods specialized for a limited scenario. In this paper, we introduce a new general algorithm based on matrix-lifting Semi-Definite Programming (SDP). The new relaxation exploits the matrix structure of the system and introduces a degradation in the performance; however, the reduction in the complexity is significant. The number of variables is decreased from $\mathcal{O}(N^2K^2)$ to $\mathcal{O}((N+K)^2)$. Moreover, this method can be implemented for any constellation and labeling method.

I. INTRODUCTION

Recently, there has been a considerable interest in Multi-Input Multi-Output (MIMO) antenna systems due to achieving a very high capacity as compared to single-antenna systems [2]. It is known that decoding is one of the important problems in MIMO systems. Decoding concerns the operation of recovering the transmitted vector from the received signal, which is known to be an NP-hard problem.

To overcome the complexity issue, a variety of sub-optimum polynomial time decoding algorithms based on Semi-Definite Programming (SDP) are suggested in the literature [1], [3]–[6].

In [3], a quasi-maximum likelihood method for decoding is introduced. Each signal constellation is expressed by its binary representation and the decoding is transformed into a quadratic minimization problem [3]. Then, the resulting problem is solved using a relaxation for rank-one matrices in SDP context. It is shown that this method has a near optimum performance and a polynomial time worst case complexity. However, the method proposed in [3] is limited to scenarios that the constellation points are expressed as a linear combination of bit labels. A typical example is the case of natural labeling in conjunction with PSK constellation [4]. Another quasi-maximum likelihood decoding method is introduced in [5] for larger PSK constellations with near ML performance and low complexity.

Another quasi-maximum likelihood decoding method is introduced in [6] for the MIMO systems employing 16-QAM. They replace any finite constellation by a polynomial constraint, e.g. if $x \in \{a, b, c\}$, then $(x-a)(x-b)(x-c) = 0$.

Then, by introducing some slack variables, the constraints are expressed in terms of quadratic polynomials. The work in [6] is restricted to MIMO systems employing 16-QAM.

In [1], an efficient approximate ML decoder for MIMO systems is developed based on SDP. The transmitted vector is expanded as a linear combination (with zero-one coefficients) of all the possible constellation points in each dimension. Using this formulation, the distance minimization in Euclidean space is expressed in terms of a binary quadratic minimization problem. The minimization of this problem is over the set of all binary rank-one matrices with column sums equal to one. In order to solve this minimization problem, two relaxation models is presented, providing a trade-off between the computational complexity and the performance (both models can be solved with polynomial-time complexity). Simulation results show that the performance of the last model is near optimal for M-ary QAM or PSK constellation (with an arbitrary binary labeling, say Gray labeling). Therefore, the decoding algorithm built on the proposed model in [1] has a near-ML performance with polynomial computational complexity.

The general method proposed in [1] has a near optimal performance for M-ary QAM or PSK constellation. However, it is more complex compared to some other methods that specialized their algorithm for a limited scenarios [1], [3]–[6]. In this paper, we introduce a new general algorithm based on matrix-lifting Semi-Definite Programming (SDP) [7], [8]. The new relaxation introduces a small degradation in the performance; however, the reduction in the complexity is significant. The number of variables is decreased from $\mathcal{O}(N^2K^2)$ to $\mathcal{O}((N+K)^2)$. Moreover, it can be implemented for any constellation and labeling method.

Following notations are used in the sequel. The space of $N \times K$ (resp. $N \times N$) real matrices is denoted by $\mathcal{M}_{N \times K}$ (resp. \mathcal{M}_N), and the space of $N \times N$ symmetric matrices is denoted by \mathcal{S}_N . For a $N \times K$ matrix $\mathbf{X} \in \mathcal{M}_{N \times K}$ the (i, j) th element is represented by x_{ij} , where $1 \leq i \leq N$, $1 \leq j \leq K$, i.e. $\mathbf{X} = [x_{ij}]$. We use $\text{trace}(\mathbf{A})$ to denote the trace of a square matrix \mathbf{A} . The space of symmetric matrices is considered with the trace inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \text{trace}(\mathbf{A}\mathbf{B})$. For $\mathbf{A}, \mathbf{B} \in \mathcal{S}_N$, $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$) denotes positive semi-definiteness (resp. positive definiteness), and $\mathbf{A} \succeq \mathbf{B}$ denotes $\mathbf{A} - \mathbf{B} \succeq 0$. For two matrices $\mathbf{A}, \mathbf{B} \in \mathcal{M}_N$, $\mathbf{A} \succeq \mathbf{B}$, $(\mathbf{A} \succ \mathbf{B})$ means $a_{ij} \geq b_{ij}$, ($a_{ij} > b_{ij}$) for all i, j . The Kronecker product of two matrices \mathbf{A} and \mathbf{B} is denoted by $\mathbf{A} \otimes \mathbf{B}$.

¹This work is supported by the Nortel Networks, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Ontario Center of Excellence (OCE).

For $\mathbf{X} \in \mathcal{M}_{N \times K}$, $\text{vec}(\mathbf{X})$ denotes the vector in \mathbb{R}^{NK} (real NK -dimensional space) that is formed from the columns of the matrix \mathbf{X} . For $\mathbf{X} \in \mathcal{M}_N$, $\text{diag}(\mathbf{X})$ is a vector of the diagonal elements of \mathbf{X} . We use $\mathbf{e}_N \in \mathbb{R}^N$ (resp. $\mathbf{0}_N \in \mathbb{R}^N$) to denote the $N \times 1$ vector of all ones (resp. all zeros), $\mathbf{E}_{N \times K} \in \mathcal{M}_{N \times K}$ to denote the matrix of all ones, and \mathbf{I}_N to denote the $N \times N$ Identity matrix. For $\mathbf{X} \in \mathcal{M}_{N \times K}$, the notation $\mathbf{X}(1:i, 1:j)$, $i < K$ and $j < N$ denotes the sub-matrix of \mathbf{X} containing the first i rows and the first j columns.

The rest of the paper is organized as follows. The problem formulation is introduced in Section II. Section III is the review of the vector-lifting semi-definite programming presented in [1]. In Section IV, we propose our new algorithm based on matrix-lifting semi-definite programming. we use the geometry of the relaxation to find a projected relaxation which has a better performance. Section V is devoted to the methods that can be used to solve the SDP problem. An augmented lagrangian method is proposed for the special structure of the problem. In Section VI, we present an optimization method, based on matrix nearness, on how we can find the integer solution of the original decoding problem from the solution of the relaxed optimization problem.

II. PROBLEM FORMULATION

A MIMO system can be modeled by [1]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where \mathbf{y} is the $M \times 1$ received vector, \mathbf{H} is $M \times N$ real channel matrix, \mathbf{n} is $N \times 1$ additive white gaussian noise vector, and \mathbf{x} is $N \times 1$ data vector whose components are selected from the set $\{s_1, \dots, s_K\}$.

Noting $x_i \in \{s_1, \dots, s_K\}$, for $i = 1, \dots, N$, we have

$$x_i = u_{i,1}s_1 + u_{i,2}s_2 + \dots + u_{i,K}s_K, \quad (2)$$

where

$$u_{i,j} \in \{0, 1\} \quad \text{and} \quad \sum_{j=1}^K u_{i,j} = 1, \quad \forall i = 1, \dots, N. \quad (3)$$

Let

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & \dots & u_{1,K} \\ u_{2,1} & \dots & u_{2,K} \\ \vdots & \ddots & \vdots \\ u_{N,1} & \dots & u_{N,K} \end{bmatrix} \quad \text{and} \quad \mathbf{s} = \begin{bmatrix} s_1 \\ \vdots \\ s_K \end{bmatrix}.$$

Therefore, the transmitted vector is $\mathbf{x} = \mathbf{U}\mathbf{s}$ and $\mathbf{U}\mathbf{e}_K = \mathbf{e}_N$.

At the receiver, the Maximum-Likelihood (ML) decoding rule is given by

$$\hat{\mathbf{x}} = \arg \min_{x_i \in \{s_1, \dots, s_K\}} \|\hat{\mathbf{y}} - \mathbf{H}\mathbf{x}\|^2, \quad (4)$$

where $\hat{\mathbf{x}}$ is the most likely input vector and $\hat{\mathbf{y}}$ is the received vector. Noting $\mathbf{x} = \mathbf{U}\mathbf{s}$, this problem is equivalent to

$$\begin{aligned} & \min_{\mathbf{U}=\mathbf{e}} \|\hat{\mathbf{y}} - \mathbf{H}\mathbf{U}\mathbf{s}\|^2 \equiv \\ & \min_{\mathbf{U}=\mathbf{e}} \mathbf{s}^T \mathbf{U}^T \mathbf{H}^T \mathbf{H} \mathbf{U} \mathbf{s} - 2\hat{\mathbf{y}}^T \mathbf{H} \mathbf{U} \mathbf{s}. \end{aligned} \quad (5)$$

Therefore, the decoding problem can be formulated as

$$\begin{aligned} \min \quad & \mathbf{s}^T \mathbf{U}^T \mathbf{H}^T \mathbf{H} \mathbf{U} \mathbf{s} - 2\hat{\mathbf{y}}^T \mathbf{H} \mathbf{U} \mathbf{s} \\ \text{s.t.} \quad & \mathbf{U}\mathbf{e}_K = \mathbf{e}_N \\ & u_{i,j} \in \{0, 1\}. \end{aligned} \quad (6)$$

Let $\mathbf{Q} = \mathbf{H}^T \mathbf{H}$, $\mathbf{S} = \mathbf{s}\mathbf{s}^T$, $\mathbf{C} = -\mathbf{s}\hat{\mathbf{y}}^T \mathbf{H}$, and let $\mathcal{E}_{N \times K}$ denote the set of all binary matrices in $\mathcal{M}_{N \times K}$ with row sums equal to one, i.e.

$$\mathcal{E}_{N \times K} = \{\mathbf{U} \in \mathcal{M}_{N \times K} : \mathbf{U}\mathbf{e}_K = \mathbf{e}_N, u_{ij} \in \{0, 1\}\}. \quad (7)$$

Therefore, the minimization problem (6) is

$$\begin{aligned} \min \quad & \text{trace}(\mathbf{S}\mathbf{U}^T \mathbf{Q} \mathbf{U} + 2\mathbf{C}\mathbf{U}) \\ \text{s.t.} \quad & \mathbf{U} \in \mathcal{E}_{N \times K} \end{aligned} \quad (8)$$

III. VECTOR-LIFTING SEMI-DEFINITE PROGRAMMING

In order to solve the optimization problem (8), the authors in [1] proposed a *quadratic vector optimization* solution by defining $\mathbf{u} = \text{vec}(\mathbf{U}^T)$, $\mathbf{U} \in \mathcal{E}_{N \times K}$. By using this notation, the objective function is replaced by $\mathbf{u}^T (\mathbf{Q} \otimes \mathbf{S}) \mathbf{u} + 2\text{vec}(\mathbf{C}^T)^T \mathbf{u}$. To solve this vector quadratic problem, the quadratic form is linearized using

$$\begin{aligned} \mathbf{Z}_{\mathbf{u}} &= \begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{u}^T \end{bmatrix} \\ &= \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix}, \end{aligned} \quad (9)$$

where $\mathbf{Y} = \mathbf{u}\mathbf{u}^T$ and it is relaxed to $\mathbf{Y} \succeq \mathbf{u}\mathbf{u}^T$, or equivalently, by the Schur complement, to the lifted constraint

$$\begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix} \succeq 0,$$

which is selected from the set

$$\mathcal{F} := \text{conv} \{ \mathbf{Z}_{\mathbf{u}} : \mathbf{u} = \text{vec}(\mathbf{U}^T), \mathbf{U} \in \mathcal{E}_{N \times K} \}. \quad (10)$$

Therefore, the decoding problem using *vector lifting semi-definite programming* can be represented by

$$\begin{aligned} \text{trace} \quad & \begin{bmatrix} 0 & \text{vec}(\mathbf{C}^T)^T \\ \text{vec}(\mathbf{C}^T) & \mathbf{Q} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix} \\ \text{s.t.} \quad & \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix} \in \mathcal{F}, \end{aligned} \quad (11)$$

which can be solved by usual SDP techniques. For more details, we refer the reader to [1].

Note that the optimization variable is a matrix in \mathcal{S}_{NK+1} . This leads to $NK + 1$ by $NK + 1$ matrix variables, which has $\mathcal{O}(N^2 K^2)$ variables and it is prohibitively large for computations. However, the best approach is to keep and exploit the structure of the original optimization problem (8).

IV. MATRIX-LIFTING SEMI-DEFINITE PROGRAMMING

In order to keep the matrix \mathbf{U} in its original form in (8), the idea is to use the constraint $\mathbf{Y} = \mathbf{U}^T \mathbf{U}$ instead of $\mathbf{Y} = \mathbf{u}\mathbf{u}^T$. Now the relaxation is $\mathbf{Y} \succeq \mathbf{U}^T \mathbf{U}$, or equivalently, by the Schur complement, $\begin{bmatrix} \mathbf{I}_N & \mathbf{U} \\ \mathbf{U}^T & \mathbf{Y} \end{bmatrix} \succeq 0$. This is known as matrix-lifting semi-definite programming.

Define the new variable $\mathbf{V} = \mathbf{U}\mathbf{S}$. Since the matrix \mathbf{S} is symmetric, the objective function in (8) can be represented as the Quadratic Matrix Program [8]

$$\begin{aligned} & \text{trace} \left(\begin{bmatrix} \mathbf{U}^T & \mathbf{V}^T \end{bmatrix} \begin{bmatrix} \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} + 2\mathbf{C}\mathbf{U} \right) \\ = & \text{trace} \left(\begin{bmatrix} \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U}^T & \mathbf{V}^T \end{bmatrix} + 2\mathbf{C}\mathbf{U} \right) \\ = & \text{trace} (\mathcal{L}\mathbf{W}_{\mathbf{U}}), \end{aligned} \quad (12)$$

where

$$\mathcal{L} = \begin{bmatrix} \mathbf{0} & \mathbf{C} & \mathbf{0} \\ \mathbf{C}^T & \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \mathbf{0} & \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix}, \quad \mathbf{W}_{\mathbf{U}} = \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{U}\mathbf{U}^T & \mathbf{U}\mathbf{V}^T \\ \mathbf{V} & \mathbf{V}\mathbf{U}^T & \mathbf{V}\mathbf{V}^T \end{bmatrix}.$$

In order to linearize $\mathbf{W}_{\mathbf{U}}$, we relax this matrix with [9]

$$\mathbf{W} \triangleq \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{X} & \mathbf{Y} \\ \mathbf{V} & \mathbf{Y} & \mathbf{Z} \end{bmatrix} \succeq 0. \quad (13)$$

On the other hand, the feasible set in (8) is the set of binary matrices in $\mathcal{M}_{N \times K}$ with row sum equal to one, the set $\mathcal{E}_{N \times K}$ in (7). By relaxing the rank-one constraint for the matrix variable in (12), we have a tractable SDP problem. The feasible set for the objective function in (12) is approximated by

$$\mathcal{F}_{\mathcal{M}} = \text{conv} \{ \mathbf{W}_{\mathbf{U}} \mid \mathbf{U} \in \mathcal{M}_{N \times K} : \mathbf{U}\mathbf{e}_K = \mathbf{e}_N, u_{ij} \in \{0, 1\}, \forall i, j; \mathbf{V} = \mathbf{U}\mathbf{S} \} \quad (14)$$

Therefore, the decoding problem can be represented by

$$\begin{aligned} \min & \quad \text{trace}(\mathcal{L}\mathbf{W}) \\ \text{s.t.} & \quad \mathbf{W} \in \mathcal{F}_{\mathcal{M}} \end{aligned} \quad (15)$$

Note that the size of matrix \mathbf{W} is $(2N + K) \times (2N + K)$, compared to $(NK + 1) \times (NK + 1)$ in [1], which is a huge reduction in the size of the problem.

Although the constraint in (13) is relaxed, we still can add/consider some linear constraints that have been removed. These constraints are valid for the non-convex rank-constrained decoding problem. However, we force the SDP problem to satisfy these constraints. Consider the auxiliary matrix \mathbf{V} and the symmetric matrices \mathbf{X}, \mathbf{Y} and \mathbf{Z} in matrix \mathbf{W} .

Since $\mathbf{U} \in \mathcal{E}_{N \times K}$ and $\sum_{j=1}^N u_{ij}^2 = 1$, it is clear that $\text{diag}(\mathbf{X}) = \mathbf{e}_N$. Also, \mathbf{Y} represents $\mathbf{U}\mathbf{S}\mathbf{U}^T$ and \mathbf{Z} represents $\mathbf{U}\mathbf{S}^2\mathbf{U}^T$. In each row of \mathbf{U} , there is only one 1 and the rest are zero. Therefore, we have

$$\text{diag}(\mathbf{Y}) = \mathbf{U}\text{diag}(\mathbf{S}) \quad \text{and} \quad \text{diag}(\mathbf{Z}) = \mathbf{U}\text{diag}(\mathbf{S}^2) \quad (16)$$

Moreover, $\mathbf{S} = \mathbf{s}\mathbf{s}^T$ (rank-one matrix) and $\mathbf{S}^2 = (\sum_{i=1}^K s_i^2)\mathbf{S}$, we have a stronger results for \mathbf{Z} , i.e. $\mathbf{Z} = (\sum_{i=1}^K s_i^2)\mathbf{Y}$. Therefore, we have

$$\begin{aligned} \min & \quad \text{trace} \left(\mathcal{L} \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{X} & \mathbf{Y} \\ \mathbf{V} & \mathbf{Y} & \mathbf{Z} \end{bmatrix} \right) \\ \text{s.t.} & \quad \mathbf{U}\mathbf{e}_K = \mathbf{e}_N ; \quad \mathbf{U} \geq 0 \\ & \quad \mathbf{V} = \mathbf{U}\mathbf{S} \\ & \quad \text{diag}(\mathbf{X}) = \mathbf{e}_N \\ & \quad \text{diag}(\mathbf{Y}) = \mathbf{U}\text{diag}(\mathbf{S}) \\ & \quad \mathbf{Z} = \left(\sum_{i=1}^K s_i^2 \right) \mathbf{Y} \\ & \quad \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{X} & \mathbf{Y} \\ \mathbf{V} & \mathbf{Y} & \mathbf{Z} \end{bmatrix} \succeq 0 \\ & \quad \mathbf{U}, \mathbf{V} \in \mathcal{M}_{N \times K}, \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathcal{S}^N \end{aligned} \quad (17)$$

A. Geometry of the Relaxation

In this section, we eliminate the constraints defining $\mathbf{U}\mathbf{e}_K = \mathbf{e}_N$ by providing a tractable representation of the linear manifold spanned by this constraint. This method is called *gradient projection* or *reduced gradient method* [10]. The following lemma is on the representation of matrices having sum of the elements in each row equal to one. This lemma is used in our reduced gradient method.

Lemma 1: Let

$$\mathbf{G} = [\mathbf{I}_{K-1} \mid -\mathbf{e}_{K-1}] \in \mathcal{M}_{(K-1) \times K} \quad (18)$$

and

$$\mathbf{F} = \frac{1}{K} (\mathbf{E}_{N \times K} - \mathbf{E}_{N \times (K-1)} \mathbf{G}) \in \mathcal{M}_{N \times K}. \quad (19)$$

A matrix $\mathbf{U} \in \mathcal{M}_{N \times K}$ with the property that the summation of its elements in each row is equal to one, i.e. $\mathbf{U}\mathbf{e}_K = \mathbf{e}_N$, can be written as

$$\mathbf{U} = \mathbf{F} + \hat{\mathbf{U}}\mathbf{G}, \quad (20)$$

where $\hat{\mathbf{U}} = \mathbf{U}(\mathbf{1} : N, \mathbf{1} : (K-1))$.

Proof: see [1]. ■

Corollary 1: $\forall \mathbf{U} \in \mathcal{E}_{N \times K}, \exists \hat{\mathbf{U}} \in \mathcal{M}_{N \times (K-1)}, \hat{u}_{ij} \in \{0, 1\}$ s.t. $\mathbf{U} = \mathbf{F} + \hat{\mathbf{U}}\mathbf{G}$, where $\hat{\mathbf{U}} = \mathbf{U}(\mathbf{1} : N, \mathbf{1} : (K-1))$. Note that the summation of each row of $\hat{\mathbf{U}}$ is 0 or 1.

Consider the minimization problem (8). By substituting (20), we can show that the decoding problem is equivalent to the following reduced matrix-lifting semi-definite programming problem:

$$\begin{aligned}
\min \quad & \text{trace} \left(\hat{\mathcal{L}} \begin{bmatrix} \mathbf{I} & \hat{\mathbf{U}}^T & \hat{\mathbf{V}}^T \\ \hat{\mathbf{U}} & \hat{\mathbf{X}} & \hat{\mathbf{Y}} \\ \hat{\mathbf{V}} & \hat{\mathbf{Y}} & \hat{\mathbf{Z}} \end{bmatrix} \right) \\
s.t. \quad & \hat{\mathbf{U}}\mathbf{e}_{\mathbf{K}-1} \leq \mathbf{e}_{\mathbf{N}} ; \hat{\mathbf{U}} \geq \mathbf{0} \\
& \hat{\mathbf{V}} = \hat{\mathbf{U}} (\mathbf{G}\mathbf{S}\mathbf{G}^T) \\
& \text{diag}(\hat{\mathbf{X}}) = \hat{\mathbf{U}}\mathbf{e}_{\mathbf{K}-1} \\
& \text{diag}(\hat{\mathbf{Y}}) = \hat{\mathbf{U}}\text{diag}(\mathbf{G}\mathbf{S}\mathbf{G}^T) \\
& \hat{\mathbf{Z}} = \left(\sum_{i=1}^{\mathbf{K}-1} (s_i - s_{\mathbf{K}})^2 \right) \hat{\mathbf{Y}} \\
& \begin{bmatrix} \mathbf{I} & \hat{\mathbf{U}}^T & \hat{\mathbf{V}}^T \\ \hat{\mathbf{U}} & \hat{\mathbf{X}} & \hat{\mathbf{Y}} \\ \hat{\mathbf{V}} & \hat{\mathbf{Y}} & \hat{\mathbf{Z}} \end{bmatrix} \succeq \mathbf{0} \\
& \hat{\mathbf{U}}, \hat{\mathbf{V}} \in \mathcal{M}_{\mathbf{N} \times (\mathbf{K}-1)}, \hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}} \in \mathcal{S}^{\mathbf{N}} \quad (21)
\end{aligned}$$

where

$$\hat{\mathcal{L}} = \begin{bmatrix} \mathbf{0} & \mathbf{G}\mathbf{S}\mathbf{F}^T\mathbf{Q} + \mathbf{G}\mathbf{C} & \mathbf{0} \\ \mathbf{Q}\mathbf{F}\mathbf{S}\mathbf{G}^T + \mathbf{C}^T\mathbf{G}^T & \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \mathbf{0} & \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix}.$$

V. SOLVING THE SDP PROBLEM

The relaxed decoding problems (17) and (21) can be solved using common interior-point methods such as SeDuMi. Using the special structure of the problem an augmented Lagrangian method based on [11] is proposed. Due to limited space it is omitted. The reader is referred to [9].

VI. INTEGER SOLUTION - MATRIX NEARNESS PROBLEM

By solving the relaxed decoding problems, we can find a solution for (17), say the matrix $\tilde{\mathbf{U}}$. In general, this matrix is not in $\mathcal{E}_{\mathbf{N} \times \mathbf{K}}$. The condition $\tilde{\mathbf{U}}\mathbf{e}_{\mathbf{K}} = \mathbf{e}_{\mathbf{N}}$ is satisfied. However, the elements are between 0 and 1. This matrix has to be converted to a 0-1 matrix by finding a nearest matrix in $\mathcal{E}_{\mathbf{N} \times \mathbf{K}}$.

A recurring problem in matrix theory is to find a structured matrix that best approximates a given matrix with respect to some distance measure. For example, it may be known a priori that a certain constraint ought to hold, and yet it fails on account of measurement errors or numerical roundoff. An attractive remedy is to replace the tainted matrix by the nearest matrix that does satisfy the constraint. Matrix approximation problems typically measure the distance between matrices with a norm. The Frobenius and spectral norms are pervasive choices because they are so analytically tractable.

In order to find the nearest solution in $\mathcal{E}_{\mathbf{N} \times \mathbf{K}}$ to $\tilde{\mathbf{U}}$, the solution of the relaxed problem, we solve the following minimization problem

$$\min_{\mathbf{U} \in \mathcal{E}_{\mathbf{N} \times \mathbf{K}}} \|\mathbf{U} - \tilde{\mathbf{U}}\|_{\mathbb{F}}^2, \quad (22)$$

where $\|\mathbf{A}\|_{\mathbb{F}}^2$ is the Frobenius norm of the matrix \mathbf{A} and is defined as $\|\mathbf{A}\|_{\mathbb{F}}^2 = \text{trace}(\mathbf{A}\mathbf{A}^T)$. Therefore, the objective function can be reformulated as

$$\begin{aligned}
\|\mathbf{U} - \tilde{\mathbf{U}}\|_{\mathbb{F}}^2 &= \text{trace} \left((\mathbf{U} - \tilde{\mathbf{U}})(\mathbf{U} - \tilde{\mathbf{U}})^T \right) \\
&= \text{trace}(\mathbf{U}\mathbf{U}^T) - 2\text{trace}(\tilde{\mathbf{U}}\mathbf{U}^T) + \text{trace}(\tilde{\mathbf{U}}\tilde{\mathbf{U}}^T) \\
&= \mathbf{N} - 2\text{trace}(\tilde{\mathbf{U}}\mathbf{U}^T) + \text{trace}(\tilde{\mathbf{U}}\tilde{\mathbf{U}}^T). \quad (23)
\end{aligned}$$

The last equality is due to the fact that for any $\mathbf{U} \in \mathcal{E}_{\mathbf{N} \times \mathbf{K}}$ we have $\text{diag}(\mathbf{U}\mathbf{U}^T) = \mathbf{e}_{\mathbf{N}}$, see (17). Therefore, after removing the constants, finding the integer solution is the answer to

$$\max_{\mathbf{U} \in \mathcal{E}_{\mathbf{N} \times \mathbf{K}}} \text{trace}(\tilde{\mathbf{U}}\mathbf{U}^T) \quad (24)$$

Consider the maximization problem

$$\begin{aligned}
\max \quad & \text{trace}(\tilde{\mathbf{U}}\mathbf{U}^T) \\
s.t. \quad & \mathbf{U}\mathbf{e}_{\mathbf{K}} = \mathbf{e}_{\mathbf{N}} \\
& \mathbf{0} \leq \mathbf{U} \leq \mathbf{1}, \quad (25)
\end{aligned}$$

where \leq in the last constraint is element-wise. This problem is a linear programming problem with linear constraints and the optimum solution is a corner point meaning that constraint are satisfied with equality at the optimum point. In other words, at the optimum point, $\mathbf{U} \in \mathcal{E}_{\mathbf{N} \times \mathbf{K}}$. Therefore, in order to find the solution for (24), we can simply solve the linear problem (25), which is strongly polynomial time.

ACKNOWLEDGEMENT

The authors would like to thank Prof. H. Wolkowicz and Yichuan Ding for helpful discussions.

REFERENCES

- [1] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, "A near maximum likelihood decoding algorithm for mimo systems based on semi-definite programming," *Accepted for publication, IEEE Trans. on Info. Theory*, April 2007.
- [2] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecomm. ETT*, vol. 10, no. 6, pp. 585-596, November 1999.
- [3] B. Steingrimsson, T. Luo, and K. M. Wong, "Soft quasi-maximum-likelihood detection for multiple-antenna wireless channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2710-2719, Nov. 2003.
- [4] B. Steingrimsson, Z. Q. Luo, and K. M. Wong, "Quasi-ML detectors with soft output and low complexity for PSK modulated MIMO channels," in *4th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2003)*, 15-18 June 2003, pp. 427 - 431.
- [5] Z. Q. Luo, X. Luo, and M. Kisiailiou, "An Efficient Quasi-Maximum Likelihood Decoder for PSK Signals," in *ICASSP '03*, 2003.
- [6] Y. C. E. A. Wiesel and S. Shamai, "Semidefinite relaxation for detection of 16-QAM signaling in mimo channels," *IEEE Signal Processing Letters*, vol. 12, no. 9, Sept. 2005.
- [7] Y. Ding and H. Wolkowicz, "A Matrix-lifting Semidefinite Relaxation for the Quadratic Assignment Problem," Department of Combinatorics & Optimization, University of Waterloo, Tech. Rep. CORR 06-22, 2006.
- [8] A. Beck, "Quadratic matrix programming," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 1224-1238, 2007.
- [9] A. Mobasher and A. K. Khandani, "Matrix-Lifting SDP for Decoding in Multiple Antenna Systems," Department of E&CE, University of Waterloo, Tech. Rep. UW-E&CE 2007-06, 2007, available via the WWW site at <http://www.cst.uwaterloo.ca/~amin>.
- [10] S. Hadley, F. Rendl, and H. Wolkowicz, "A new lower bound via projection for the quadratic assignment problem," *Math. Oper. Res.*, vol. 17, no. 3, pp. 727-739, 1992.
- [11] S. Burer and R. D. C. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Mathematical Programming (Series B)*, vol. 95, pp. 329-357, 2003.